A Robust and efficient stereo matching algorithm

Ruihua Ma, Monique Thonnat

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A robust and efficient stereo matching algorithm

Ruihua MA
Monique THONNAT

N° 1860
Février 1993

PROGRAMME 4

Robotique, Image
et
Vision

Rapport de recherche

1993

A Robust and Efficient Stereo Matching Algorithm

Ruihua MA and Monique THONNAT

INRIA Sophia-Antipolis
B.P.93, 06902 Sophia-Antipolis Cedex FRANCE
Tel: (33) 93-65-78-66 Fax: (33) 93-65-77-66
email: ruihua@sophia.inria.fr, thonnat@sophia.inria.fr

Abstract

Matching is recognized as the most difficult step in stereo vision. Many algorithms have been proposed but they are not robust enough. The reason for this, in short, is that the available matching constraints are not effectively used. In fact, the design and the implementation of a stereo algorithm are closely coupled issues. To obtain a robust algorithm, constraints should be appropriately used. In addition, noise in the data must also be taken into account explicitly and effectively. Therefore, in this paper, we discuss thoroughly the stereo matching problem, beginning by a brief discussion on the choice of matching primitives. Then matching constraints are discussed in detail and their use justified, especially on the disparity continuity and the choice of a concrete measure of this constraint. Our main contributions are the following. (1) By making realistic assumptions, we show that a disparity gradient (DG) limit as small as 0.2 can be used. Support from neighboring points is not weighted in order to make full use of structural (shape) information, which is made possible by the small DG limit. (2) Noise (due to discretization and dislocation) is fully taken into account and effectively dealt with in the implementation. This in turn makes the use of a small DG limit possible. (3) Disparity discontinuity along an edge chain can be dealt with by requiring only piecewise continuity. In fact, it is frequently observed that a edge chain results from different surfaces due to the edge detection and linking. Abundant tests have been performed, mostly on complex real scenes. Three of them are shown and analyzed in this paper. The new algorithm proves to be robust and efficient.

Key words: stereo matching, disparity continuity, small disparity gradient limit, shape information, discretization noise, edge dislocation.

Un Algorithme Robuste et Efficace pour la Mise en Correspondance Stéréoscopique

Résumé

Le problème de mise en correspondance (*matching*) est l'étape la plus difficile en stéréo. Beaucoup d'algorithmes de matching ont été proposés mais rares sont ceux qui sont robustes. La raison en est, en un mot, que les contraintes de matching disponibles ne sont pas exploîtées de manière efficace. En effet, la conception d'un algorithme stéréo et son implémentation sont des problèmes étroitement liés. Afin d'obtenir un algorithme robuste, non seulement les contraintes doivenent être utilisées de manière appropriée, mais le bruit de données doit être pris en compte explicitement et effectivement. Ainsi ce papier se propose de donner une analyse complète du problème de stéréo. Nous commençons par une brève discussion sur le choix de primitives de matching. Ensuite nous discutons comment utiliser des contraintes, en particulier celle de la continuité de disparité et comment choisir une mesure concrète pour cette contrainte. Un algorithme est proposé en basant sur ces analyses. Nos contributions sont les suivantes. (1) En adoptant une approche probabiliste et des hypothèses réalistes, nous démontrons que une limite de gradient de disparité de 0.2 peut être utilisée. (2) Les bruits dans les données sont explicitement prises en compte, concernant en particulier celui dans la position des primitives. (3) Le chaînage bruité est également traité. Par chaînage bruité, nous entendons l'existance de chaînes de contours ayant des éléments provenant de plusieures surfaces dans l'espace. De nombreux tests ont été effectués, dont trois d'entre eux sont analysés dans ce papier. Ce qui permet d'affirmer la robustesse et l'efficacité de l'algorithme.

Mots clés: matching stéréo, continuité de disparité, petite limite de gradient de disparité, information de forme, bruit de dislocalisation.

Contents

1 Introduction

The recovery of 3-D structure of scene is one of the fundamental problems in computer vision. The use of stereopsis as a passive technique for this purpose has been widely investigated over the past years. While the underlying principle – triangulation – is straightforward, the task of identifying (matching) homologous items in two (or more) images has proved to be very difficult[3, 4]. (This problem is commonly referred to as *correspondence problem*). Research in this area has been conducted in two directions. Theoretical work attempted to model human (or, more generally, biological) stereopsis in order to provide a general computational framework[17, 16, 18]; whereas most efforts have been made in the development of stereo algorithms for practical applications[4, 8].

Marr considered the stereo as a low-level module[16], which is supported by Julesz's famous random-dot stereogram experiment[16]. In their computational theory, Marr and Poggio[17] proposed two constraints for the correspondence problem, uniqueness and continuity, both based on physical observations. Mayhew and Frisby[18] introduced another constraint called "figural continuity". These constraints govern the interaction between good matches, that is, global consistency. Global constraints, together with local ones such as similarity and epipolar geometry, constitute the base and starting point for the design of stereo algorithms.

However, theoretical works do not show us how the available constraints should be explored in an algorithm, neither how they should be implemented in a computer program. Rather, one is free in making his own choices in the design and implementation. Concretely, one must consider the following problems in the design of a stereo algorithm: *What primitives are to be used? How should the continuity constraint be used? Admitting the existence of a disparity gradient (DG) limit, what value should be chosen? How to cope with noise in the data? How to choose an adequate compromise between conflicting requirements?* etc..

Different answers to these questions lead to different algorithms. The fact that many stereo algorithms have been proposed and new ones are still emerging comes from the fact that existing algorithms are not robust enough. The reason for this, in short, is that the available matching constraints are not effectively used. So in this paper, we first make a thorough discussion on these problems, based on which we propose a new algorithm. Our main contributions are the following:

1. By making realistic assumptions, we show that a DG limit as small as 0.2 can be used (appendix C). Support from neighboring points is not weighted in order to make full use of structural (shape) information, which is made possible by the small DG limit (§ 4).

2. Noise (due to discretization and dislocation) is fully taken into account and effectively dealt with in the implementation. This in turn makes the use of a small DG limit possible (§ 3 and § 4).

3. Disparity discontinuity along an edge chain can be dealt with by requiring only piecewise continuity. In fact, it is frequently observed that a edge chain results from different surfaces due to the edge detection and linking (§ 4).

The remainder of this paper is organized as follows. We first discuss briefly matching primitives (§ 2). Then we give a detailed analysis of the matching problem (§ 3), focusing on the points listed above. Based on this analysis, our new algorithm is described (§ 4). After experimental results are reported (§ 5), a general performance evaluation will be given (§ 6). Finally, we conclude with some perspectives (§ 7).

2 What to match?

At sight, this question would appear inappropriate, because the choice of matching primitives should depend mainly on the imaging system and the application. However, this choice must also take into account practical considerations, i.e., the matching primitives should be general, available, and matchable.

Generality means that primitives should convey the majority of useful information in any given scene. *Availability* means there must be techniques to extract them reliably. Finally, *matchability* means the there must be way to apply effectively matching constraints.

Many types primitives are found in the literature. They vary from low-level ones such simple intensity pixels, edges, intermediate-level ones such contours (made up of linked edges), line segments, curves, corners, regions, to high-level ones such as structured forms. But rare are those satisfying all of the three requirements mentioned above. Roughly speaking, lower level primitives can be extracted more easily the high level ones. Primitives easy to obtain are difficult to match, and *vice versa*. For example, regions would be ideal for surface recovering. But region segmentation itself proves to be a very difficult problem. No currently available segmentation algorithm available yields satisfactory results. Intensity pixels are widely used in area-based algorithms[13, 10]. However, apart from the sensitivity to perspective distortion, these methods operating intensity often fail at occluding boundaries of surfaces, because such locations cannot be coped with correctly unless they are already detected, which again require a non-trivial prior segmentation. This is particularly unacceptable considering the fact that a great deal of information about a scene is found at such locations. In fact the disparity continuity constraint, which should allow to select good matches or suppress false ones, cannot be exploited appropriately[13, 14] (In [10], it is even completely ignored). Line segments may be interesting to describe indoor scenes, but are not suitable to scenes containing complex forms as found in outdoor situations. Further the extraction of line segments suffers from spurious merging and splitting.

Edges represent sharp variations in image intensity. In general they describe pertinent events in a scene such as object boundaries, region marks, occluding boundaries

of surfaces, or shadows of objects in 3-D space. They can be extracted easily using efficient edge detectors such as those described in [6, 7]. Further, by careful use of available matching constraints, edges can be matched reliably. Particularly, the disparity continuity constraint can be appropriately applied. We add also that linked edges are ready to be fitted to curves or line segments, allowing thus both indoor and outdoor scenes to be described.

In short, we have seen that matching primitives and matching methods are related issues. Among all types of primitives, edges seem to be the most attractive. In the algorithm presented below, linked edges are used as input data. But this does not necessarily mean that edges should be the only type of matching primitives. Rather, we can think of algorithms working on different primitives which operate in a cooperative way (see § 7).

3 The Matching process

Matching is a search. To carry it out, constraints are needed to find a unique solution. In stereo matching, as noted by Marr ([16], p.115), the available constraints are usually *necessary*, yet *not* necessarily *sufficient*. Furthermore, the stereo matching is not a mathematically well formulated problem.

In an ill-constrained search, as is the case for stereo matching, the following observations are relevant, when the search space increases:

- the probability of error increases;
- the computational cost also adds up, often in a more rapid manner.

Stereo matching in general, suffers primarily from the first problem, referred to as the *false target problem*[17]. To reduce the search space, one first restricts the regions where corresponding primitives are looked for. Then the properties of these primitives are examined: only primitives with similar properties are retained as candidates. At this stage, *local* constraints are sufficient yet the solution is not unique.

To approach the solution, candidate matches are submitted to global consistency checking using some additional constraints. The *global* constraints may or may not be sufficient. The underlying idea is that while good matches tend to be mutually consistent (in terms of disparity continuity, *e.g.*), this is generally not the case for false ones. Therefore global constraints can be used for two purposes: (1) match selection; (2) false match detection and elimination. Notice that in practice the two operations are usually performed separately.

Thus two central problems in resolving the stereo matching are:

- restriction of search space;

- efficient use of global constraints,

We note that these considerations are to a certain extent in conflict with the requirement to match as many as possible image primitives. Thus compromises must be made. In addition, one must take into account the noisy and discrete nature of the input, which is decisive in the design of a stereo matching algorithm. This will be highlighted along the discussions below.

3.1 Local constraints

Local constraints include epipolar geometry and similarity. They are intended to limit the number of candidate matches for a given primitive in one image.

Epipolar geometry

The epipolar geometry allows to restrict the search in one dimension, i.e., on lines called epipolar lines. For non-parallel stereo setups, since image coordinates are discrete values whereas an epipolar lines are calculated in a continuous space, the search region consists of points immediately above and under the epipolar lines, i.e., it is a "ribbon" of 2 pixel width, rather than a simple straight line (Figure 1).

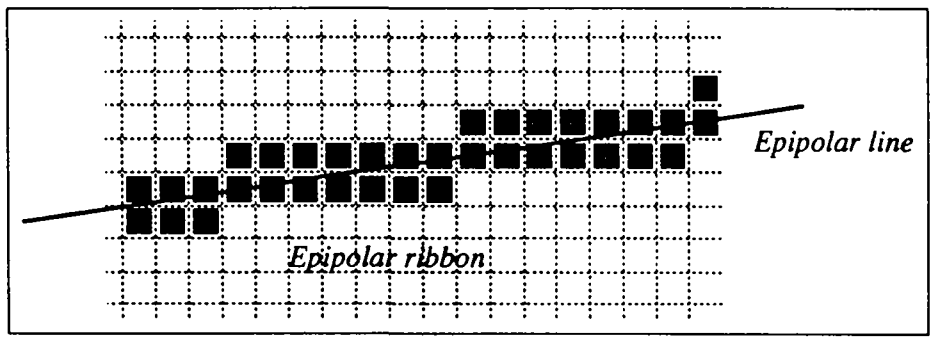

Figure 1: Epipolar line in a discrete image.

Similarity

Except for extreme cases, surface patches in a scene projected into two cameras form similar image patches. This local similarity information is primary and has been extensively used in searching. In case of edges, gradient magnitude and orientation can be used. However the following problems inhibit the excessive emphasis on the use of this information:

- the presence of noise;
- the perspective distortion, which produces the deformed image patches and eventually dissimilar patches at occluding boundaries.

Thus a strict similarity criterion is meaningless, especially at occluding boundaries. In other words, retained candidates should not be discriminated in the subsequent steps for deciding good matches.

3.2 Global constraints

Global constraints are derived from physical properties of scenes. Marr and Poggio [17] made the following relevant observations (1) a given point on physical surface has a unique position in space at any one time; (2) matter is cohesive and separated into objects. Furthermore, the surfaces of objects are generally such that the changes in surface are very small compared with their distance from the viewer.

Marr and Poggio [17] proposed two constraints based on their observations:

1. *Uniqueness.* Each item from each image may be assigned at most one disparity value.

2. *Continuity.* Disparity varies smoothly almost every where over the image.

These constraints are rather general and not ready for use unless some conditions are satisfied. This is the case for the disparity continuity constraint in area-based algorithms.

Mayhew and Frisby [18] introduced the notion of *figural continuity* which states that in an image a chain of connected edge points probably represents the projection of a continuous curve in 3-D space. This figural continuity is directly justified by Marr and Poggio's second observation and suggests the disparity continuity along any edge chain. This constraint can be readily applied to edge chains, yet a concrete measure of continuity remains to be determined. Also care should be taken for the case where an edge chain results from different surfaces. In such a case, the disparity is at best piecewise continuous.

3.2.1 Disparity continuity constraint

Given two points in the space and their projections in the two image planes, Burt and Julesz [5] define disparity gradient (DG) as the ratio of the difference of disparity to the separation of their cyclopean images for binocular fusion. They also showed evidence of the existence of a DG limit in human vision system.

Inspired by the experience of Burt and Julesz, Pollard, Mayhew and Frisby propose to impose a DG limit for the disambiguation purpose in their algorithm [22]. They analyzed in detail the significance of imposing such a constraint, whereas their efforts to prove the existence of a DG limit of 1 (corresponding to human vision ability) does

not seem convincing and meaningful. In fact, in [26], a proof is given that a DG limit of 2 exists provided that scene surfaces are opaque.

It has been reported in [22, 23] that small DG limit provides better disambiguating power. It would be more straightforward to say that a small DG limit yields more reliable matches, or allows smaller error percentage. The authors also pointed out that the choice of a DG limit should be a compromise between disambiguating power and generality – the ability to cope with as wide a range of surfaces as possible. We propose to use a small DG limit. By small, we mean a DG limit much less than 0.5. The choice of DG limit will be discussed in detail in § 3.4.

3.2.2 Uniqueness

Marr and Poggio's uniqueness constraint cannot be immediately applied to image points (or edges). In fact, a pixel has physical size and each pixel results from a small portion of a surface. It is not difficult to imagine a situation in which a same portion of a sloping surface in a scene form a certain number of pixels in one image, but more (or less) in another one.

When a small DG limit is used, perspective distortion, if any, will not be significant. Moreover with the epipolar geometry constraint, the worst case in applying the uniqueness constraint on edges is that some points among connected edges find no match, yet others do. In case where connected edges result from sloping surfaces and lie on an epipolar line, it is generally out of the question to match them unambiguously. Notice that for the first case, an interpolation is feasible to fill the holes (unmatched points) by making use of their matched neighbors and the figural continuity.

3.2.3 Ordering

This constraint states that the relative positional order of image primitives along an epipolar line in one image, is preserved in the other image. The constraint is a conditional one because of the existence of the so-called *forbidden zone*[27]. When the opacity assumption is verified, the ordering constraint is valid for most real cases and should be useful for stereo matching.

From an implementation point of view, the following remarks can be made. First, to use it, the stereo pair should be rectified. Second, when integrated in the matching process, as in [2] and [19], computational efficiency does not seem satisfactory. Finally if this constraint is used to detect matches which violate it, it is not possible to decide from which match the violation arises. In [23] a solution based on the sum of "matching strength" is proposed but not justified. In short, the ordering constraint is, for us, somewhat restrictive, either in its use or in computational efficiency.

3.3 Coarse-to-fine control structure

The *coarse-to-fine* matching strategy was originally proposed as part of Marr and Poggio's human vision theory [17]. However, we are attracted more by its practical merit in the sense that it allows to limit the search space and thus to reduce greatly the number of false matches. The idea is to guide the search of potential matches using matches established at a coarser level, which should be reliable. A supplementary benefit is the reduction of computational cost (see [14] for an analysis).

One has to note that false matches at coarser levels may lead to false matches at their subsequent levels. But this is, in general, much less significant compared to errors made by methods matching directly at full resolution, at least with the results reported and to our knowledge.

3.4 Geometric constraints: determining appropriate values

Following the work by Arnold and Binford[1], Stewart gave a more detailed analysis on the derivation of geometric constraints in stereo[24, 25][1]. Based on the assumption that line segments in space are distributed uniformly on a Gaussian sphere, some probability density functions (PDF's) are derived, among which are the PDF of orientation difference between corresponding line segments and the PDF of the disparity gradient.

We are concerned with the determination of the threshold on orientation difference and the DG limit. As one can see, small thresholds on these parameters allow less candidate matches to be retained, i.e., the disambiguation process can be eased. But how small can they be, precisely? We show that a threshold of $\pi/6$ on orientation difference (appendix B) and a DG limit as small as 0.2 (appendix C) can be used, allowing more than 90% of surfaces to be treated. This is done by making use of Stewart's results and some realistic assumptions. Notice that all derivations are realized in continuous domain, care should be taken in dealing with noise due to image discretization and dislocation in edge detection (see § 3.4.1).

In [22, 23], a DG limit of 0.5 has been proposed, which will enable 92.5% of scene surfaces to be coped with at the relative depth $\bar{z} = 5$ (see appendix A). Since a limit of 0.2 is sufficient to deal with the majority of surfaces, a DG limit greater than 0.2 can be said to be a sacrifice to the overall matching performance knowing that small DG limit allows a better disambiguating power.

3.4.1 Implementation problem

Let a chosen DG limit to be DG_0. Imposing such a DG limit is to say

$$\frac{|\Delta disp|}{dist} \leq DG_0$$

[1] This does not mean that we agree with all the points presented in these papers.

where $\Delta disp$ denotes the disparity difference between a pair of points and *dist* their image distance. The inequality above can be rewritten as

$$|\Delta disp| \leq DG_0 \cdot dist = \Delta disp_0(dist) \quad (1)$$

Thus the DG limit is translated into an allowed disparity difference: $\Delta disp_0(dist) = DG_0 \cdot dist$. The computation involves image coordinates and the latter suffer from discretization error. In the case of edge features, there exists also dislocation errors. These errors should be taken into account. This is done by adding a constant Δd to the allowed disparity difference, i.e., $\Delta disp_0(dist) = DG_0 \cdot dist + \Delta d$. In practice, Δd is in the order of 1 2 pixels.

The effect of the term Δd is illustrated by Figure 2. In the figure are shown two

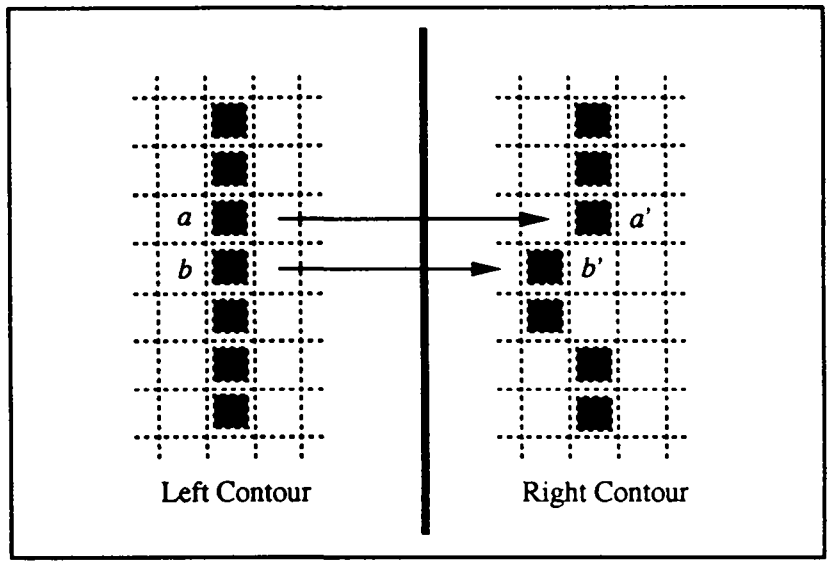

Figure 2: Edge dislocation in discrete image.

corresponding chains. However, two elements of the right chain are not lined up as it should be because of dislocation in the edge detection. Obviously, the two matches (a, a') and (b, b') should be consistent on the criterion of disparity continuity. Without the term Δd, a $DG_0 < 1$ would fail to consider (a, a') and (b, b') to be consistent because $dist_{aa'} = 1$ and thus $\Delta disp_0(dist) < 1$. However, by adding the constant term Δd, this case can be quite easily and naturally treated. Moreover, doing this does not imply sacrificing the disambiguating power, for a small DG limit can still be used.

In closing this section, we point out that a same constraint (*e.g.*, DG limit) can be implemented quite differently in different algorithms. However, it does not make sense to tune parameters derived in the continuous domain for ideal cases without being able to cope with noise properly. This explains in part why different algorithms using similar constraints perform quite differently.

4 The Algorithm

In this section, we describe an algorithm based on the discussions in the previous sections[2]. The matching constraints implemented in this algorithm are: epipolar geometry, similarity, disparity continuity, and uniqueness. The coarse-to-fine strategy is adopted: each image of the original stereo pair is consolidated to form a pyramid of images representing different resolution levels. Feature extraction is carried out independently for each level.

4.1 A voting scheme for matching

Our algorithm is basically a voting scheme, as illustrated in Figure 3, in which points on the chains A and B in the left image are supposed to match points on the chains C and D. We proceed the matching point by point and epipolar lines are supposed to be parallel to the horizontal axis of the image planes. The matching is divided into two steps: searching candidates and choosing matches.

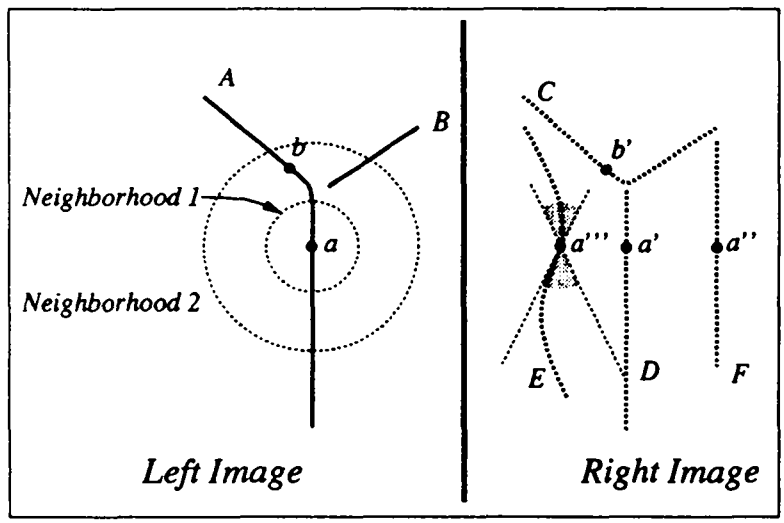

Figure 3: A voting scheme for matching.

In the first step, for each edge point in the left image, candidate matches are searched in the right image. The search is performed by applying a similarity criterion on the epipolar lines within a predefined interval. Thus each point in the left image may have more than one candidates, or no candidate at all. For example, in the figure, the point a may have three candidates: a', a'' and a'''.

[2]An early presentation can be found in [9], where the stereo algorithms developed at INRIA are presented and compared.

In the second step, candidate matches of point are submitted to a global consistency checking to determine a match. Consider the point a in the figure. First, we define a neighborhood of a, N_2 for example. Now let us take the candidate (a, a') and look at the candidates of all neighbors of a. The neighbor b has one candidate (b, b'). Obviously the pair (a, a') and (b, b') will give rise to a disparity gradient near zero, thus less than a predefined DG limit DG_0. So we call this "b vote for (a, a') or (a, a') receives a vote (termed as *support* in the algorithm. See next section) from b". But (a, a'') won't receive vote from b because the pair (a, a'') and (b, b') yields a disparity gradient bigger than DG_0. The same happens to (a, a'''). Similarly, (a, a') will receive vote from almost every other neighbor. Both (a, a'') and (a, a''') may receive some votes, for example, from some neighbors on the vertical part of the chain A, but not from all points in N_2. Thus the candidate (a, a'), which receives more votes from its neighbors than (a, a'') and (a, a'''), is retained as a good match. The same operation is applied to every point in the left image.

Several remarks are to be made here. First, if a smaller neighborhood is adopted, N_1 for example, then it may happen that (a, a') and (a, a'') even (a, a''') will have the same number of votes. In such a case, we will not be able to decide the best match from the candidates. This is in fact a case reduced to a repeating pattern. Therefore, taking an appropriate (not too small) neighborhood size is critical for the disambiguation. This is rather natural due to the fact that bigger neighborhood conveys more structural information.

Second, the predefined DG limit DG_0 does affect the disambiguating process. The bigger DG_0, the bigger the allowed disparity difference (see (1)). The allowed disparity difference increases with the distance. It will not be surprising to see that the candidate match (a, a''') will receive as many votes as (a, a') if DG_0 is big (> 0.5, *e.g.*), which make the disambiguation impossible. However, if DG_0 is small (0.2, *e.g.*) it will receive less votes and (a, a') receives the most votes. This is why a small DG limit is preferred for the disambiguation.

Third, this voting scheme has no restriction on the way the corresponding points are linked into chains in the two images. As in Figure 3, the points on the chains A and B in the left image will be matched to those on C and D without any problem. This difference in configuration of points affects only the neighborhood for any given point. As long as a small DG limit and a reasonably big neighborhood are adopted, the result of matching remains the same. This feature is important because primitive extraction is basically a low-level operation that does not guarantee to yield edge chains with exactly the same configuration.

Fourth, as one may notice, the votes are not weighted, which is made possible by using a small DG limit (See justification in the next section).

4.2 Implementation

The algorithm is divided into the following steps: (1) candidate searching; (2) score computation; (3) match validation; (4) noise suppression; (5) selection of matches among candidates; (6) edge chain fractioning; (7) disparity interpolation. The same operations are applied to all resolution levels.

The first time the match validation is done, it may happen that some matchable points are still unprocessed. The score computation and match validation can be iterated in order to match more points.

Candidate searching

Given a point to be matched, the search takes place along the epipolar line associated with the point within a disparity range, either estimated (at the coarsest level) or predicted (at finer levels). Edges encountered in the search region are then submitted to similarity examination. This involves G and θ, the gradient norm and gradient orientation, respectively. Those satisfying the following conditions are retained as candidates:

$$|G - G'| \leq \Delta G_0$$

$$|\theta - \theta'| \leq \Delta \theta_0$$

where ΔG_0 is the threshold for the difference in gradient norms; and $\Delta \theta_0$ that for the difference in gradient orientations.

Computing score

This operation can be thought of as counting the votes a candidate match receives, just as described in the precedent section. But the terms "vote" is replaced by "support" and "number of votes" by "score", since some modification and refinement are under way.

Suppose we have a edge chain $\mathcal{C}$ composed of L ordered points: $\mathcal{C} = \{p_i, i \in [1, L]\}$, each point having a list of candidate matches: $\mathcal{M}_i = \{m_{ij}, j \in [1, n_i]\}$, where n_i the number of candidates of the ith point p_i. Thus the score of the candidate m_{ij}, defined as the sum of support it receives, is formulated as follows:

$$Score\ (m_{ij}) = \sum_{e_k \in \mathcal{N}(e_i)} max\{Support\ (m_{ij}|m_{kl}), m_{kl} \in \mathcal{M}_k\} \qquad (2)$$

where $\mathcal{N}(e_i)$ denotes e_i's neighborhood and $Support\ (m_{ij}|m_{kl})$ the support m_{ij} receives from m_{kl} of e_k.

$\mathcal{N}(e_i)$ is an interval defined on $\mathcal{C}$ around e_i, expressed as $[N_1, N_2]$. Given the predefined radius of neighborhood NS, we have

$$N_1 = \begin{cases} 1, & \text{if } i - NS < 1; \\ i - NS, & \text{otherwise} \end{cases} \text{ and } N_2 = \begin{cases} L, & \text{if } i + NS > L; \\ i + NS & \text{otherwise} \end{cases}$$

The determination of *Support* $(m_{ij}|m_{kl})$ is shown in Table 1: where $dist(e_i, e_k)$ is

Table 1: Support determination.

	Support $(m_{ij}\|m_{kl})$	
$\|disp(m_{ij}) - disp(m_{kl})\|$	m_{kl} is a candidate	m_{kl} is a validated match
$= 0$	2	4
$< dist(e_i, e_k) \cdot DG_0 + 1$	1	2
$> dist(e_i, e_k) \cdot DG_0 + 1$	0	0

the city distance between e_i and e_k, i.e., $|\Delta u(e_i, e_k)| + |\Delta v(e_i, e_k)|$.

Several notes are in order. First, city distance is used for its simplicity. In the worst case,

$$|\Delta u(e_i, e_k)| = |\Delta v(e_i, e_k)|,$$

$$|\Delta u(e_i, e_k)| + |\Delta v(e_i, e_k)| = \sqrt{2} \cdot \sqrt{\Delta u(e_i, e_k)^2 + \Delta v(e_i, e_k)^2},$$

DG_0 can be thought to be multiplied by $\sqrt{2}$. It won't be significant for a small DG_0. However, the computation is greatly reduced. In addition, we build look-up table for the allowed disparity difference, using the fact that the "distance" involves only integers.

Second, zero disparity difference is favored in support determination (Table 1). In case where more than one candidates of e_k can gives support, the maximal support is taken into account (see (2)).

Third, the support is not weighted by $dist(e_i, e_k)$, in contrast to most other algorithms such as [20]. This is very important. It is clear that weighting support by distance penalize support from distant neighbors and the structural information cannot be fully used. Indeed, weighting was intended to reduce the influence of noisy candidates, especially when a relatively big DG limit is used. But one must know that weighting support only makes the disambiguating process more sensitive to noise, for it implies using small neighborhood. Our algorithm imposes strict constraint on candidates that give support by using a small DG limit. The perturbation of noisy candidates can thus be reduced to a very low level.

Fourth, once retained, candidate matches are considered "equal". Similarity is not involved in the score computation.

Fifth, in principle, the bigger the neighborhood radius NS is, the better the matching result. But bigger NS gives rise to increased computational cost. An alternative

allowing the use of big NS without increasing the computation cost consists in sampling neighbors, using every 2 or 3 neighbors for the score computation.

Finally, score computation is by nature a parallel process.

The score computation is performed in both images of the stereo pair.

Validating matches

As in [20, 11], matches are validated in a reciprocal manner. Suppose a point e in the left image and a point e' in the right image are involved in a candidate match, i.e., $\{e, e'\}$ is in the candidate list of e and $\{e', e\}$ that of e'. If both $\{e, e'\}$ and $\{e', e\}$ have the maximal score, the two candidates are validated. As can be expected, this manner of validation allows to greatly reduce the number of false matches. Matched points are no more considered as candidate of any unmatched point, in order to impose the uniqueness constraint.

Noise suppression

Noisy matches are those which violate the figural continuity and DG limit constraints, translated into disparity continuity in the algorithm. A validated match of a point is suppressed if it is not consistent with validated matches of the neighboring points at either side of the point. Here the disparity continuity is required in a single-sided way.

Selection of matches among candidates

After some iterations of the score computation and match validation, there are unmatched edge points which have matched neighboring points and hold still a candidate list. We pick up as matches those candidates satisfying the DG limit. This operation has the same purpose as interpolation to fill-in holes (unmatched points) on a chain, but matches obtained in this way are guaranteed to correspond to physically existing points. The disadvantage is the increased computational cost compared to the interpolation.

Edge chain fractioning

Edge chains are cut at locations where the disparity is continuous at only one side. This can be considered as part of depth segmentation [15]. Moreover, the operation prevents the disparity interpolation from being performed between edges underlying disparity discontinuity.

Disparity interpolation

Once noisy matches are eliminated and chains with discontinuous disparity fractioned, the interpolation of disparity can be performed with safety for unmatched points in a chain.

5 Experiments

The algorithm has been tested on numerous stereo pairs, including indoor, outdoor and synthetic scenes. By doing so, we hope to evaluate the performance of the algorithm, in particular its robustness, a critical measure for all stereo algorithms.

The three scenes we show here are chosen for their complexity. The depth variation in the scenes is important, permitting to see whether occluding boundaries can be well dealt with.

Evaluation of stereo algorithms based on results is always a tough task due to lack of well-defined procedures, criteria, and database with ground truth. So here we give both numerical results and images for qualitative analysis. Different views of the results of 3-D reconstruction are also provided to allow a visual evaluation. Analysis, qualitative and/or quantitative, will be made as well, whenever possible.

In this section, we first report the parameters used in the experiments. Then the results are presented and analyzed.

5.1 Input data

The stereo imaging system uses two CCD cameras and is calibrated.

The input consists of essentially chains made up of connected edges. Edges are detected by the edge detector described in [7]. Edge linking is performed by using the algorithm presented in [12] and the chains serve as support for global consistency checking . Each edge point is associated with its magnitude and orientation.

5.2 Parameter set

We summarize the parameter set involved in the algorithm.

1. *Similarity criteria*: ΔG_0 and $\Delta\theta_0$.
   They are both fixed to $\pi/6$ in our experiments. Note that raw images have 256 gray levels.

2. *Neighborhood size*: NS.
   It is fixed to 30 pixels, corresponding to a neighborhood of $[-30, 30]$.

3. *Radius of search region* for estimated disparity: $[-3, 3]$, in pixel.

4. *Disparity gradient limit*: DG_0. We have fixed it to 0.2.

5. *Width of the epipolar ribbon*: fixed to 2 pixels.

6. *Disparity range.*
   They are variable and in a calibrated stereo system are supposed to be known. The values we use are found in Table 2.

7. *Number of iterations* for the score computation and match validation.
   It is fixed to 3.

During the whole test, only the disparity range is changed accordingly. The rest of the parameter set remains the same for all stereo pairs. Although all parameters are modifiable, they affect little on the matching results, provided that these lie in reasonable intervals. The failures reported result rather from the characteristics of the images (*e.g.*, repeating patterns) than from inappropriate choice of certain parameters.

5.3 Results

Synthetic scene

This scene contains a cylinder, a cone, an ashtray, a torus, and a calibration grid. The surfaces of the first three objects are textured. From Figure 4, we can see that almost all edge points visible to both views are matched. The 3-D visualization[3] shows that no false match is found. Zigzags in the reconstructed scene are due to the discrete effect in edge position. Notice that textures (including squares on the grid in a large sense) have been correctly matched. This is because the directions of repetition do not coincide with that of epipolar lines, rather than that our algorithm is able to deal with repeating patterns (see § 6).

Rocks and grid

This scene contains rocks and a calibration grid on the ground in front of a wall. Due to the small depth, the perspective distortion is important, especially for the rocks. This explains why only 55% (=9372/17047) of edge points find their potential matches (see Table 2). Still, edges resulting from main structures of the scene are well detected on both images and matched.

Nonetheless, the matching suffers from the repeating squares on the grid. Only a portion of edge points of these are matched and this owing to the coarse-to-fine control

[3]Reconstructed points of a same chain are linked. This same is true for all 3-D visualizations.

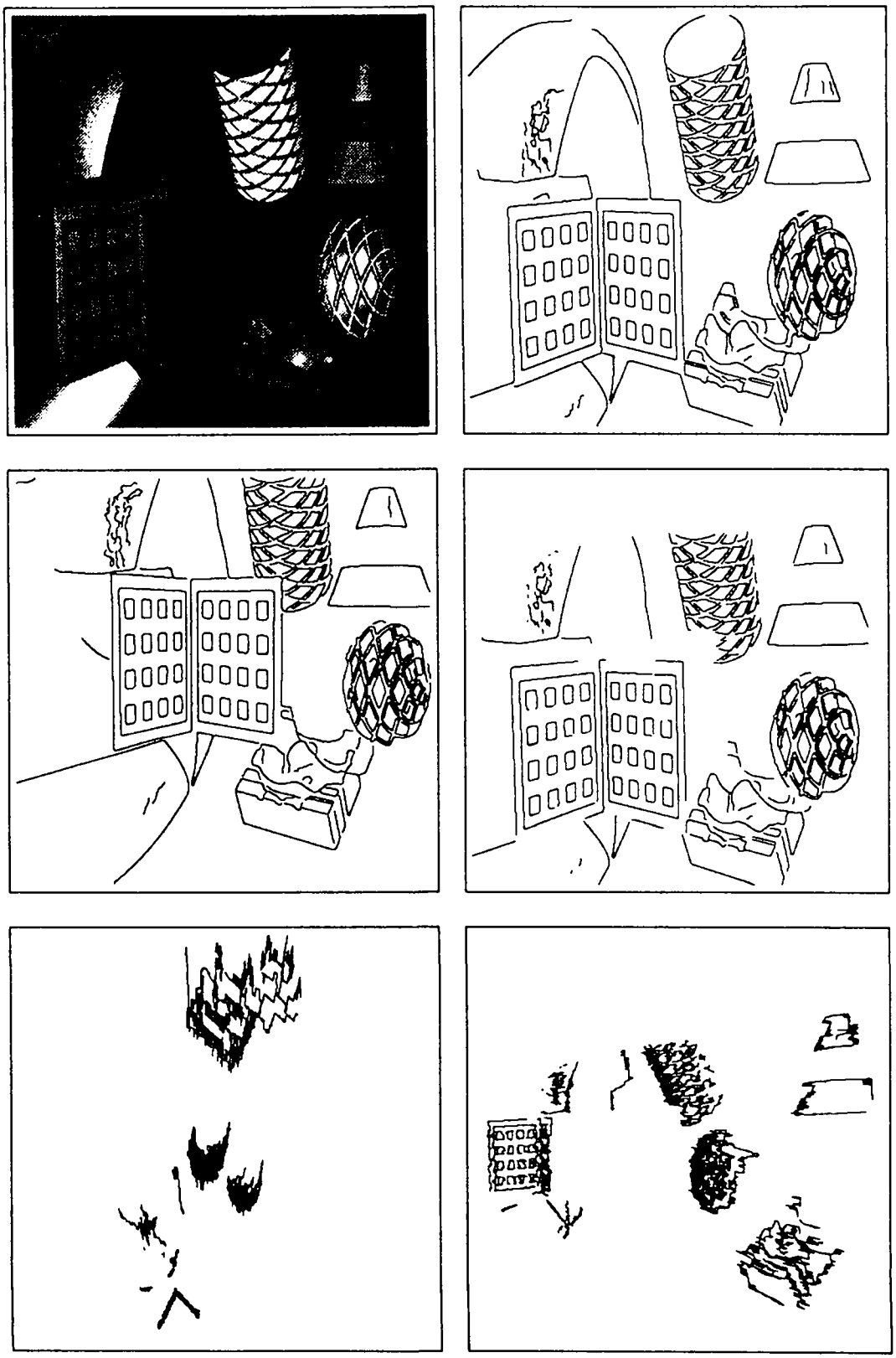

Figure 4: Synthetic scene. (a) Original image (View1). (b) Edge image (View 1). (c) Edge image (View 2). (d) Matched and reconstructed edges (View 1). Reconstructed scene: (e) Top view. (f) Side view. (The images are arranged in left-to-right and top-to-down order. The same is true for other scenes.)

structure. Part of the chains on the top of the grid are not well matched, probably due to two factors: (1) they lie on the epipolar lines; (2) the calibration and thus the computation of epipolar lines are not of good quality. This is also why zigzags in the reconstructed scene are more noticeable than those in the synthetic scene.

Road

This is a complex road scene composed of moving objects (cars, van, cyclist), road markings, trees, etc.. Edges are well matched. The line indicated by an arrow in Figure 6(f) results from the false edge points on the right side of the images.

The numerical results are resumed in Table 2. The CPU time is measured on a Sun SPARCworkstation IPC.

Table 2: Numerical results.

Scene	Disp range	Nb of points		Nb of matches					CPU time (s)
		T	C	V	N	P	I	F	
Synthetic	[0, 70]	14374	12074	9990	15	649	1548	12143	94.3
Rocks and grid	[50, 97]	17047	9372	7451	68	475	2834	10598	115.0
Road	[-40, 10]	18265	12278	10510	44	541	2304	13204	97.9

T: total number of edge points in the left image (or View 1; the same for the rest);

C: number of points having at least one candidate match;

V: number of matches obtained by the match validation procedure;

N: number of noisy matches among V (suppressed);

P: number of matches picked up from candidate lists and satisfying the disparity continuity constraint;

I: number of matches obtained by disparity interpolation;

F: final total number of matches;

From Table 2 we can know that over 80% of points having candidate matches are validated within 3 iterations, with about 80% in the first iteration. Among final matches, up to 25% are obtained by interpolation. The chain fractioning guarantees the safety of this operation.

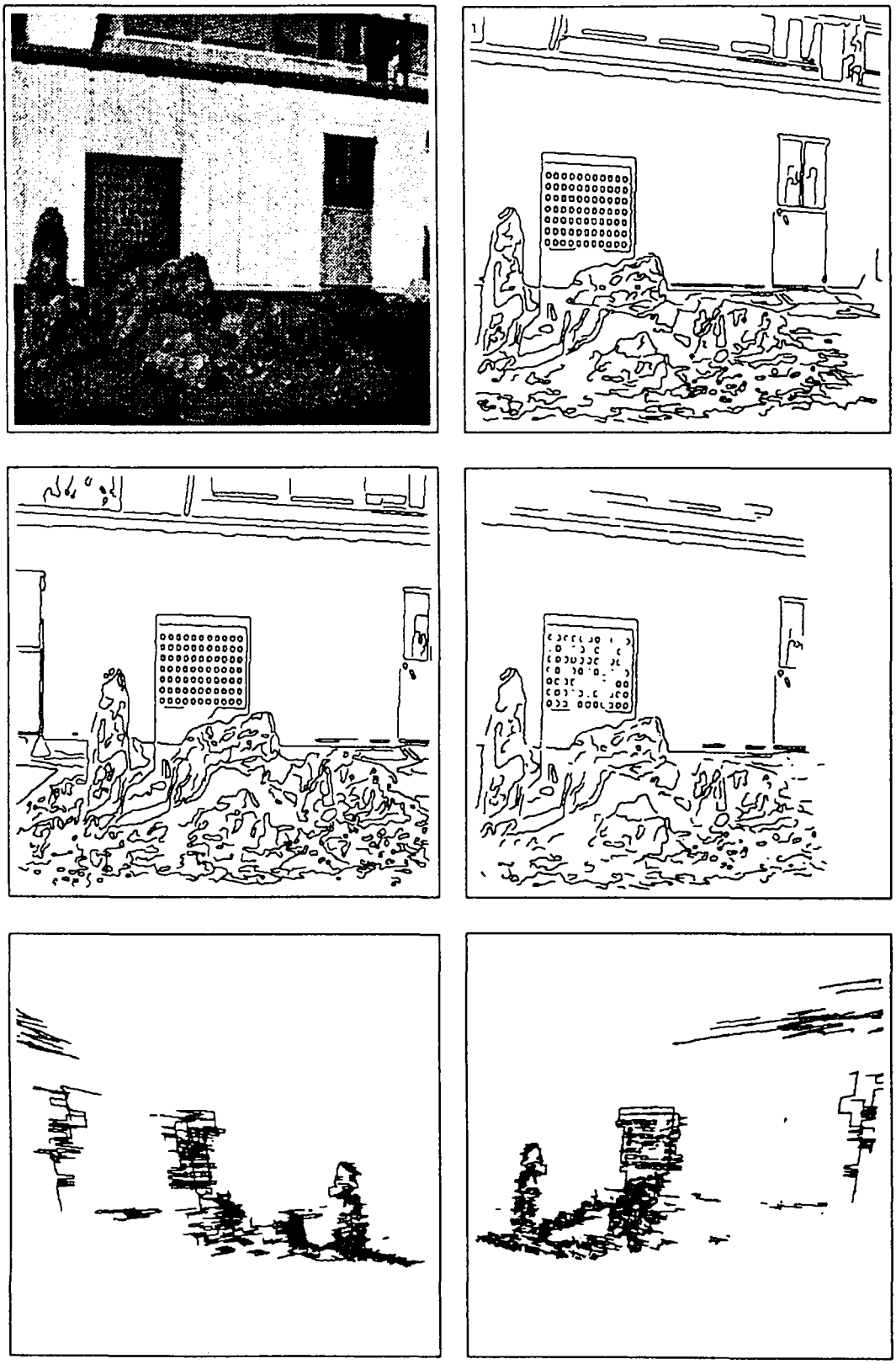

Figure 5: Rocks and grid. (a) Original image (View1). (b) Edge image (View 1). (c) Edge image (View 2). (d) Matched and reconstructed edges (View 1). Reconstructed scene: (e) Leftside view. (f) Rightside view.

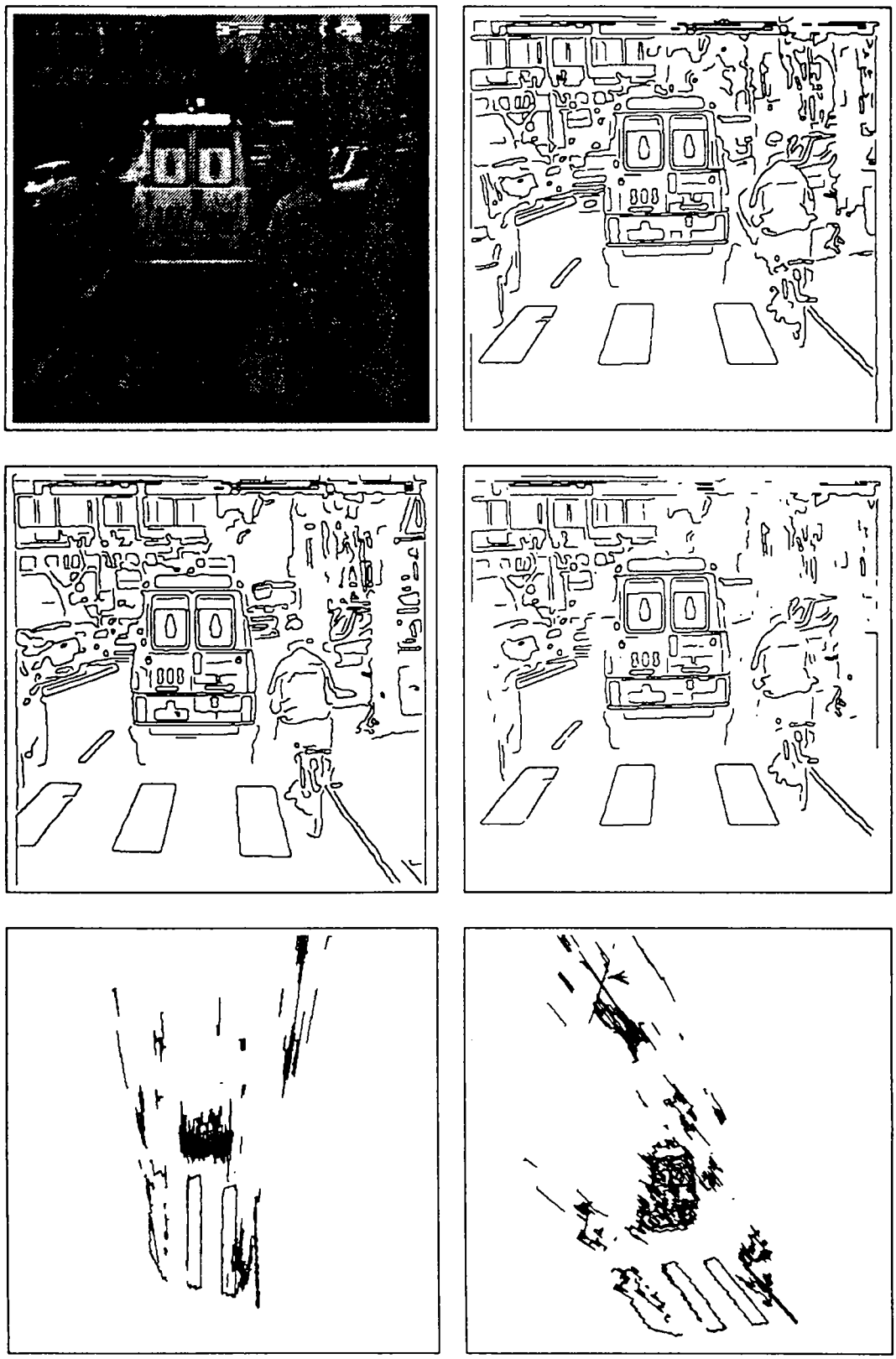

Figure 6: Road scene. (a) Original image (Left). (b) Edge image (Left). (c) Edge image (Right). (d) Matched and reconstructed edges (Left). Reconstructed scene: (e) Top view. (f) Left top view.

6 Evaluation of the general performance

Apart from the scenes presented in this paper, we have tested our algorithm on the JISCT[4] Stereo Evaluation Package[5] which includes 45 stereo pairs (or triplets). Based on all our experiments, we have self-evaluated the performance of the algorithm in general. The result is shown here.

We focus our attention on the following aspects: (1) robustness; (2) CPU time; (3) failures and ill-fitted cases.

6.1 Robustness

Robustness has always been our primary concern during the development of this algorithm.

Except the problems discussed in the next section, the robustness of the algorithm is evidenced by the following points:

1. The parameters in the algorithm are practically insensitive to different types of scenes.

2. The use of DG limit allows to cope with stereo images which have undergone relatively significant perspective distortions.

3. By relying more on the global consistency constraint and only imposing a rough similarity criterion, occluding contours can be well matched.

4. The algorithm is not restrictive vis-a-vis the shape of chains. In fact, a group of points lying on a same chain, as long as they come up to form candidate matches with their counter parts in the other image, will be maximally scored, regardless how the counterparts are grouped, i.e., the shape of chains these counterparts form. This can also be seen in Eq. (2).

5. Except for cases the algorithm is unable to cope with (§ 6.3), the majority of matchable points are matched.

6. The error percentage is quite small, owing to the use of the small DG limit which is made possible by the constant term taking into account the dislocation of edges in discrete images.

[4] Abbreviation for JPL, INRIA, SRI International, CMU (Carnegie-Mellon Univ.), and Teleos.

[5] In the context of the JISCT Stereo Evaluation process aiming at the evaluation of the effectiveness of existing stereo algorithms.

6.2 CPU time

The CPU time is divided into two parts: that spent on the extraction of primitives and that on the matching. The part for the matching in general dominates the part of extraction primitives, despite of the additional time needed to obtain edge images at coarser resolutions.

In matching, the score computation is the most time-consuming. We observe that among validated matches, over 70% result from the first iteration. So in principle, there is no question if the number of iteration is set to a minimal number, say, 2. The rest of the matching can be completed by the following steps – selecting matches among candidates and interpolation.

As discussed in § 4.2, an alternative to reduce further the time for the scoring is to use some kind of sampling technique. Also in § 4.2, we mentioned that the score computation is parallelizable. So to further speed up the matching process, special hardware and/or parallel architectures can be used.

The overall execution time does not appear to be excessive, but rather short. We indicate also that no special effort has been done to optimize the source code.

6.3 Failures and ill-fitted cases

Failures are found principally with repeating patterns. Here we consider chain segments parallel to the scanlines as a special case of repeating patterns, with the edge point being the smallest element.

As to the inability in the presence of significant perspective distortion, we think that the problem lies rather in what to match than in how to match. In fact, in such cases, corresponding features are not available because of the dissimilarity of image patches. However, we believe that in most application contexts, a stereo system should be arranged in such a way that perspective distortion, if present, is reasonably small.

Another failure case concerns scenes some portions of which do not contain sufficient salient features or markings (including apparent occluding contours), as in the rocks scene.

In summary, our algorithm fails in two typical cases: the one related to repeating patterns, the other to the lack of salient features. The first problem is inherent to binocular stereo and is still open; the second is a proper problem of the algorithm (*cf.* area-based methods work well on textured regions, which do not necessarily have salient features).

7 Conclusion and perspectives

In this paper, we have made a thorough analysis of stereo vision, concentrating on the correspondence problem. In particular, the importance of the ability of coping with noise has been emphasized. This has led to a simple edge chain based stereo algorithm in which noise is dealt with explicitly and efficiently. The algorithm is fully tested on various scenes and it turns out to be fast and robust in normal circumstances.

The result obtained with this algorithm might have insufficient richness of the disparity information. However, since the algorithm yields reliable matches, we think it is particularly promising to combine this algorithm with other schemes such as area correlation based ones, which suffer from occluding contours.

Acknowledgment

The synthetic scene is due to Zhengyou Zhang and the 3-D visualization program due to Luc Robert, Robotvis Project, INRIA Sophia-Antipolis.

A Parallel setup and relative depth

Technically speaking, a parallel setup is hard to obtain. However, there are two advantages in considering such a system. First, mathematical relations are simpler. Second, results derived from it remain valid, *qualitatively*, for other non-parallel setups. To us, qualitative remarks are more important because noise in data and non-parallel setups make it impossible to use results derived derived for a parallel setup in continuous domain.

In a parallel setup of stereo system (Fig. 7), there exists an inverse proportionality relation between the disparity and the depth:

$$d = u_2 - u_1 = -\frac{B\,f}{z}$$

where f is the focal length of the cameras and B the baseline length.

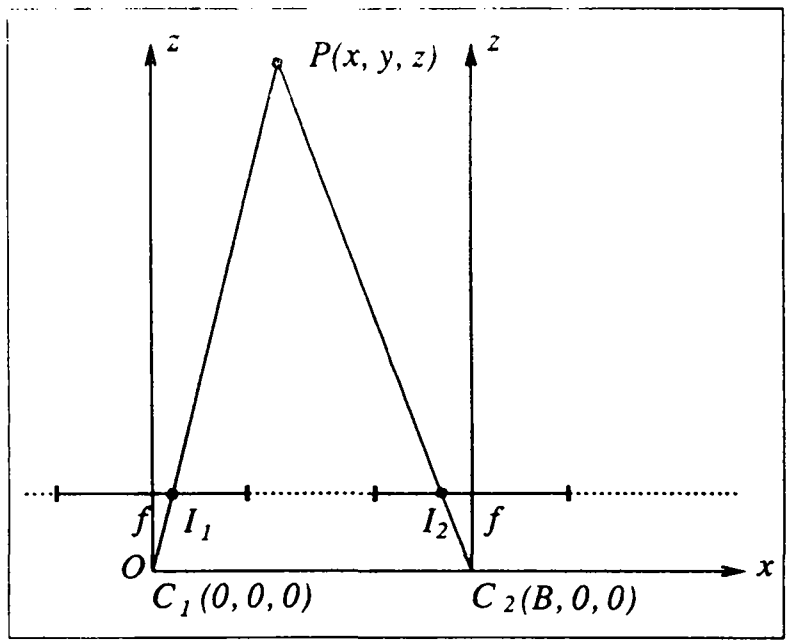

Figure 7: A parallel setup of stereo system.

When discussing the disparity, it is more convenient to consider the relation between d and z/B. It is also more relevant, since when talking about the distance, the notion of nearness of farness is related to the baseline length of the system. We call $\bar{z} = z/B$ *relative depth*. Thus d is only related to $\bar{z}$ and all derivations remain valid no matter what the baseline length B is. In the following, the term *depth* should be referred to $\bar{z}$.

In real applications, we have often a minimal depth $\bar{z}_{min}$ in mind. Scene events within $\bar{z}_{min}$ will not be considered. The reason for this is twofold: (1) the common view field of the two cameras can be reasonably big beyond this depth; (2) the region

for candidate searching is also limited. Assuming $\bar{z}_{min} = 5$ is, if not too reserved, realistic. For example, if $B = 0.4m$, $\bar{z}_{min} = 5$ corresponds to a minimal depth of $2m$. As to $\bar{z}_{max}$, it changes from scene to scene and may also be subjective, as for $\bar{z}_{min}$.

B Threshold on orientation difference

Stewart derives in [24] a PDF (an approximative form) as follows:

$$p_\alpha(\alpha_L, \alpha_R, \bar{z}) = \frac{d^2 f \sin\alpha_L}{2\pi \sin^2\alpha_R [f^2\zeta^2 + d^2]^{3/2}}, 0 < \alpha_L, \alpha_R < \pi,$$

where $\zeta = \cos\alpha_L - \sin\alpha_L \cot\alpha_R$, α_L and α_R are the orientations of the projections of a line segment in space in the left and right images, respectively.

Using relative depth and simplifying ζ into $\zeta = \sin(\alpha_R - \alpha_L)/\sin\alpha_R$, we get

$$p_\alpha(\alpha_L, \alpha_R, \bar{z}) = \frac{\bar{z}\sin\alpha_L \sin\alpha_R}{2\pi[\sin^2\alpha_R + \bar{z}^2 \sin^2(\alpha_R - \alpha_L)]^{3/2}} \quad (3)$$

This equation shows clearly that α_L and α_R tend to be equal, which agrees with our intuition. Further, for a given orientation difference, the probability decreases with the depth $\bar{z}$. This function is illustrated[6] in Figure 8, where $\bar{z}$ is fixed to $\bar{z}_{min}$ (= 5).

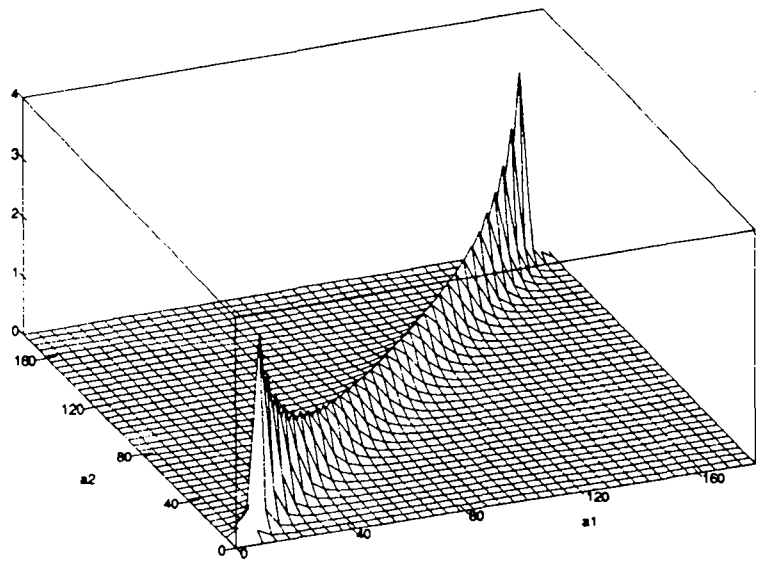

Figure 8: The PDF of the orientation of the projections (α_L and α_R in degrees.

[6]This is also given in [1], but not the explicit equation.

It can be seen that the flattest part is around $\{\alpha_L = \pi, \alpha_R = \pi\}$. In other words, the largest orientation difference should be found in the case where the projection of a line segment in one image is vertical. Take an extreme case, say, for $\alpha_L = \pi$. The probability that the orientation difference is less then $\Delta\alpha$ will be

$$P(|\alpha_L - \alpha_R| < \Delta\alpha, \bar{z}) = \frac{\int_{\pi-\Delta\alpha}^{\pi+\Delta\alpha} p_\alpha(\pi, \alpha_R, \bar{z}) d\alpha_R}{\int_0^{\pi} p_\alpha(\pi, \alpha_R, \bar{z}) d\alpha_R} \quad (4)$$

Fixing a $\Delta\alpha$, this probability decreases when the depth $\bar{z}$ increases. For example, if $\Delta\alpha = \pi/6$, then $P(|\alpha_L - \alpha_R| < \Delta\alpha, 5) = 94.5\%$ and $P(|\alpha_L - \alpha_R| < \Delta\alpha, 10) = 99.9\%$. If we consider $\bar{z}_{min} = 5$, the probability 94.5% above represents the worst case, both for projection orientations and for the depth. This means that a threshold $\Delta\alpha$ of $\pi/6$ is big enough to treat the majority of edge features in images.

Although the above example also suggests that given a probability of the allowed orientation difference, more difference should be allowed when the depth increases (i.e., a varying threshold), we think that $\pi/6$ is plausible, for the following reasons. First, in real cases, the system setup is in general non-parallel. The result above only gives us a qualitative idea of the order of the probability, but not an exact value to rely on. To us, it suffices to know that it is safe to use such a threshold. Second, the threshold cannot be too small because of the noise in orientation measures. Third, it is simpler to use a fixed threshold than a dynamic one.

C The disparity gradient limit

Stewart derives another PDF (also an approximate form) concerning the disparity gradient as follows in [24]:

$$p_{DG}(DG, \bar{z}) = \frac{\bar{z} \cdot DG}{2(\bar{z}^2 \cdot DG^2 + 1)^{3/2}}, -\infty < DG < \infty$$

or

$$p_{DG}(DG, \bar{z}) = \frac{\bar{z} \cdot DG}{(1 + \bar{z}^2 \cdot DG^2)^{3/2}}, 0 < DG < \infty \quad (5)$$

where DG is the disparity gradient, defined as the ratio of disparity difference of the endpoints of a line segment over the length of the segment in one image[7].

Our problem is to find a DG limit which allows the majority of surfaces in a scene to be treated. Here the "majority" means "> 90%". In [21], the authors have considered the cumulative probability at different depths, expressed by the following equation:

$$P(DG < DG_0) = \int_0^{DG_0} p_{DG}(DG, \bar{z}) dDG$$

[7] Although this definition is different from that using cyclopean distance as in [21], it is not surprising to see that the derivations are the same. This is due to approximations made during the derivation.

$$= \frac{\bar{z} \cdot DG_0}{\sqrt{1 + \bar{z}^2 \cdot DG_0^2}} \tag{6}$$

This function is drawn in Figure 9.

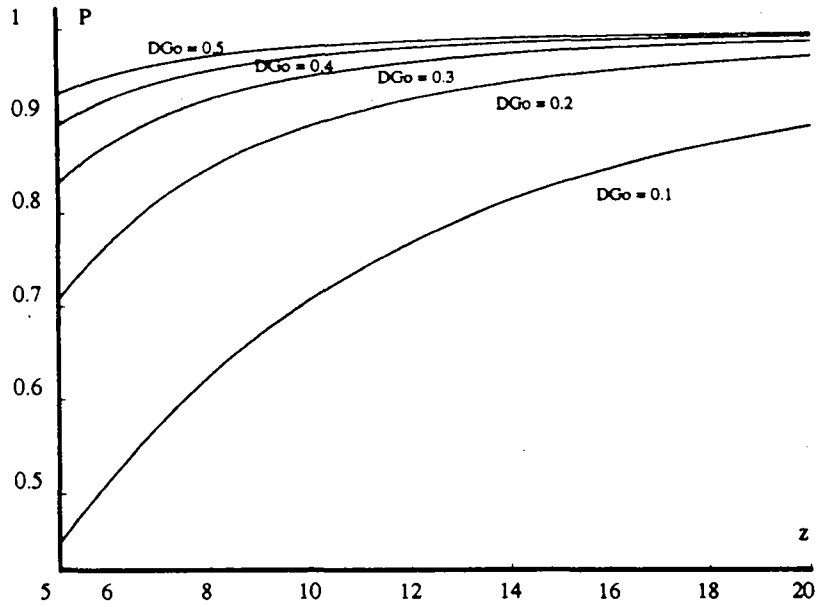

Figure 9: The cumulative probability on a DG limit.

From (6), we can see that at $\bar{z} = 5$, over 70% of the scene surfaces can be correctly treated using a DG limit of 0.2, whereas about 30% of them will present a DG beyond this limit and not be dealt with correctly. However, surfaces resulting in DG beyond 0.2, when projected into images, represent in general only a small part of the entire images, because of their (deep) sloping nature with respect to the image plane. It makes sense to consider that image patches resulted from these surfaces cover less than 30% of the entire image. In such cases, only less than 10% of image features may not be correctly handled, using a DG limit of 0.2. In other words, at $\bar{z} = 5$, $DG_0 = 0.2$ allows more than 90% of image primitives to be correctly dealt with. It is worth noting that (1) the ability of a DG limit to deal with sloping surfaces increases with the depth, (2) all objects in a scene are not at $\bar{z}_{min}$ but range from $\bar{z}_{min}$ to $\bar{z}_{max}$, and (3) for deep sloping surfaces, some may not be visible to both cameras, or the dissimilarity is so large that corresponding scene features cannot be detected in both images with available techniques. These observations make us more confident about the use of a DG limit of 0.2.

References

[1] R.D. Arnold and T.O. Binford. Geometric constraints in stereo vision. In *Proceedings of SPIE Image Processing for Missile Guidance*, volume 238, pages 281–292, San Diego, 1980.

[2] H.H. Baker and T.O. Binford. Depth from edge and intensity based stereo. In *Proc. 7th Intl. Joint Conf. Artificial Intell.*, pages 631–636, Vancouver, Canada, August 81.

[3] D.H. Ballard and C.M. Brown. *Computer Vision.* Prentice-Hall, Englewood Cliffs, NJ, 1982.

[4] S.T. Barnard and M.A. Fischler. Computational stereo. *ACM Computing Surveys*, 14(4):553–572, December 1982.

[5] P. Burt and B. Julesz. Modifications of the classical notion of Panum's fusional area. *Perception*, 9:671–682, 1980.

[6] J. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on PAMI*, 8(6):679–698, 1986.

[7] R. Deriche. Using Canny's Criteria to Derive a Recursively Implemented Optimal Edge Detector. *International Journal of Computer Vision*, 1(2):167–187, 1987.

[8] U.R. Dhond and J.K. Aggarwal. Structure from stereo – a review. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1489–1510, November 1989.

[9] O. Faugeras, P. Fua, B. Hotz, R. Ma, L. Robert, M. Thonnat, and Z. Zhang. Quantitative and qualitative comparision of some area- and feature-based stereo algorithms. In *Proceedings of the International Workshop on Robust Computer Vision*, Bonn, Germany, March 1992.

[10] P. Fua. A parallel stereo algorithm that produces dense depth maps and preserves image features. Research report RR 1369, INRIA, Fance, January 1991.

[11] P. Fua. A parallel stereo algorithm that produces dnse depth maps and preserves image features. *Machine Vision and Applications*, 1991. Accepted for publication, also available as INRIA research report 1369.

[12] G. Giraudon. An Efficient Edge Following Algorithm. In *Proceedings of 5th Scandinavian Conference on Image Analysis*, pages 547–554, Stockholm, 1987.

[13] W.E.L. Grimson. *From Images to Surfaces: A Computational Study of the Human Early Visual System.* M.I.T. Press, Cambridge, MA, 1981.

[14] W.E.L. Grimson. Computational experiments with a feature based stereo algorithm. *IEEE Transactions on PAMI*, 7(1):17–34, January 1985.

[15] R. Ma and M. Thonnat. Object detection in outdoor scenes by disparity map segmentation. In *Proceedings of the 11th International Conference on Pattern Recognition*, volume A, pages 546–549, The Hague, The Netherlands, 1992.

[16] D. Marr. *Vision*. W.H. Freeman, New York, 1982.

[17] D. Marr and T. Poggio. A theory of human stereo vision. Technical Report AI Memo No. 451, MIT, November 1977.

[18] J.E.W. Mayhew and J.P. Frisby. Psychophysical and computational studies toward a theory of human stereopsis. *Artificial Intelligence*, 17:349–385, 1981.

[19] Y. Ohta and T. Kanade. Stereo by intra- and inter-scanline search. *IEEE Transactions on Pattern Anal. and Machine Intell.*, 7(2):139–154, March 1985.

[20] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. Disparity gradient, lipschitz continuity, and computing binocular correspondance. In *Proceedings of the 3rd International Symposium of Robotics Research*, pages 19–26, Gouvieux, France, 1985.

[21] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. Disparity gradient, lipschitz continuity, and computing binocular correspondance. Technical Report AIVRU010, AI Vision Research Unit, University of Sheffield, 1985.

[22] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. PMF: A stereo correspondence algorithm using a disparity gradient limit. *Perception*, 14:449–470, 1985.

[23] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. Implementation details of the PMF stereo algorithm. In J.E.W. Mayhew and J.P. Frisby, editors, *3D Model Recognition From Stereoscopic Cues*, chapter 2, pages 33–39. The MIT Press, 1990.

[24] C.V. Stewart. On the derivation of geometric constraints in stereo. Technical Report 91-26, Department of Computer Science, Rensselaer Polytechnique Institute, Troy, New York, August 1991.

[25] C.V. Stewart. On the derivation of geometric constraints in stereo. In *Proceedings of the Conf. on Computer Vision and Pattern Recognition*, pages 769–772, 1992.

[26] H.P. Trivedi and S. A. Lloyd. The role of disparity gradient in stereo vision. In J.E.W. Mayhew and J.P. Frisby, editors, *3D Model Recognition From Stereoscopic Cues*, chapter 2, pages 47–50. The MIT Press, 1990.

[27] A.L. Yuille and T. Poggio. A generalized ordering constraint for stereo correspondance. Technical Report A.I. Memo 777, MIT, May 1984.

Unité de Recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 SOPHIA ANTIPOLIS Cedex (France)

Unité de Recherche INRIA Lorraine Technopôle de Nancy-Brabois - Campus Scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 VILLERS LES NANCY Cedex (France)
Unité de Recherche INRIA Rennes IRISA, Campus Universitaire de Beaulieu 35042 RENNES Cedex (France)
Unité de Recherche INRIA Rhône-Alpes 46, avenue Félix Viallet - 38031 GRENOBLE Cedex (France)
Unité de Recherche INRIA Rocquencourt Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

EDITEUR
INRIA - Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

ISSN 0249 - 6399